Morphologically constrained and data informed cell segmentation of budding yeast

Morphologically Constrained and Data Informed Cell Segmentation of Budding Yeast

Elco Bakker [1,2], Peter S. Swain [1,2+], and Matthew M. Crane [1,2+]

November 24, 2017

Abstract

1 Motivation:

Although high-content image cytometry is becoming increasingly routine, processing the large amount of data acquired during time-lapse experiments remains a challenge. The majority of approaches for automated single-cell segmentation focus on flat, uniform fields of view covered with a single layer of cells. In the increasingly popular microfluidic devices that trap individual cells for long term imaging, these conditions are not met. Consequently, most techniques for segmentation perform poorly. Although potentially constraining the generalizability of software, incorporating information about the microfluidic features, flow of media and the morphology of the cells can substantially improve performance.

2 Results:

Here we present DISCO (Data Informed Segmentation of Cell Objects), a framework for using the physical constraints imposed by microfluidic traps, the shape based morphological constraints of budding yeast and temporal information about cell growth and motion to allow tracking and segmentation of cells in microfluidic devices. Using manually curated data sets, we demonstrate substantial improvements in both tracking and segmentation when compared with existing software.

3 Availability:

The MATLAB® code for the algorithm and for measuring performance is available at `https://github.com/pswain/segmentation-software` and the test images and the curated ground-truth results used for comparing the algorithms are available at `http://datashare.is.ed.ac.uk/handle/10283/2002`.

4 Introduction

One of the primary methods through which information is acquired from biological samples is by optical imaging. Imaging by both transmitted light and fluorescence is essential to modern biology, and the proliferation of innovative imaging techniques continues to increase its importance ([1]). The automated application of these imaging methodologies, often in time-lapse microscopy experiments, has left biomedical researchers with a deluge of data, and a common bottleneck to analysis is the necessary segmentation into either cells or other regions of interest.

This challenge, recognized for nearly fifty years, has been the subject of intense research ([2]). The most widely used and most generalizable methods rely on thresholding images into a foreground and background ([3]). Nevertheless, these methods have several problems that warrant the development of bespoke tools ([4]). Importantly, a fluorescent marker is often used to label either part of the cell ([5, 6, 7, 3, 8, 9, 10]) or the media ([10]) to achieve an acceptable accuracy of segmentation. For example, a large proportion of the ORF-GFP library in budding yeast ([11]), over four thousand cell lines, were tagged with red fluorescent protein to facilitate automated segmentation ([12]). Such

markers can both increase the workload of constructing strains and interfere with other measurements of fluorescence ([13]) and always occupy a fluorescence channel, limiting the amount of data that can be acquired. Fluorescent proteins can provide substantial information about the biology of the cell, and thus requiring fluorescence channels specifically for segmentation can preclude desirable experiments. Additionally, fluorescence imaging damages cells and should be kept to a minimum ([14]).

There is thus a need for reliable segmentation based on bright-field and differential interference contrast (DIC) images. Existing approaches mostly assume a uniform field of view and therefore can preclude images obtained in microfluidic devices, which are rapidly being adopted to study single cells ([15]). Many of these approaches rely primarily on local image features ([16, 17, 10, 18]) or a combination of fluorescence and bright-field features ([19]).

Methods for image segmentation, and computer vision in general, have to balance trade-offs between generalizability and precision. This requirement is especially acute in imaging in the life sciences, where a wide range of model organisms and imaging environments are employed ([3, 20, 21, 6]). Methodologies that apply to all these diverse organisms and experimental conditions are necessarily agnostic about the constraints that are specific to a particular case. With this limitation in mind, we here confined our interest to the automated segmentation of cells of *S. cerevisiae* in microfluidic experiments: specifically, long-term imaging using devices containing traps for individual cells (Fig. 1). The microfluidic device we predominantly consider is ALCATRAS ([22]), with additional images provided using the device developed by [23].

Widespread interest in the replicative aging of single cells has resulted in an explosion in the number of microfluidic devices that can trap mother cells of the eukaryote *Saccharomyces cerevisiae* for their entire lifespan ([24, 25, 26, 27, 22, 28]). Budding yeast divide rapidly—growing exponentially in glucose with a doubling period of 80-90 minutes. To image the same cells over a long period of time, newborn cells (daughters) must be removed to prevent the device from becoming overcrowded. In contrast to typical tracking, where there is only a small probability of losing a tracked cell if the cell either dies or moves outside the field of view ([29]), this removal means that cells regularly appear, disappear, and replace each other.

Here we present a comprehensive framework to segment and track budding yeast cells. By focusing on budding yeast in microfluidic traps, we can leverage prior knowledge about shape, motion and appearance to improve accuracy and performance. This approach, employing both fitted probabilistic models and supervised machine learning, is generally applicable and can provide substantial improvements in accuracy. An example of its benefits is a study of the response of yeast to osmotic stress where precise volume and signalling dynamics of single cells were obtained ([30]).

5 Approach

Our framework for integrated identification, segmentation and tracking of cells is structured into four stages:

1. the microfluidic features of the traps are located and used to define regions of interest and to register images;

2. a classifier based on supervised learning is applied to each pixel to score whether the pixel is interior to the cell, an edge, or background;

3. seeds for cellular locations at the current time point are proposed based on both the classification of pixels and the location of cells at previous time points;

4. a shape-based active contour is iteratively applied to the proposed seeds until the image is segmented.

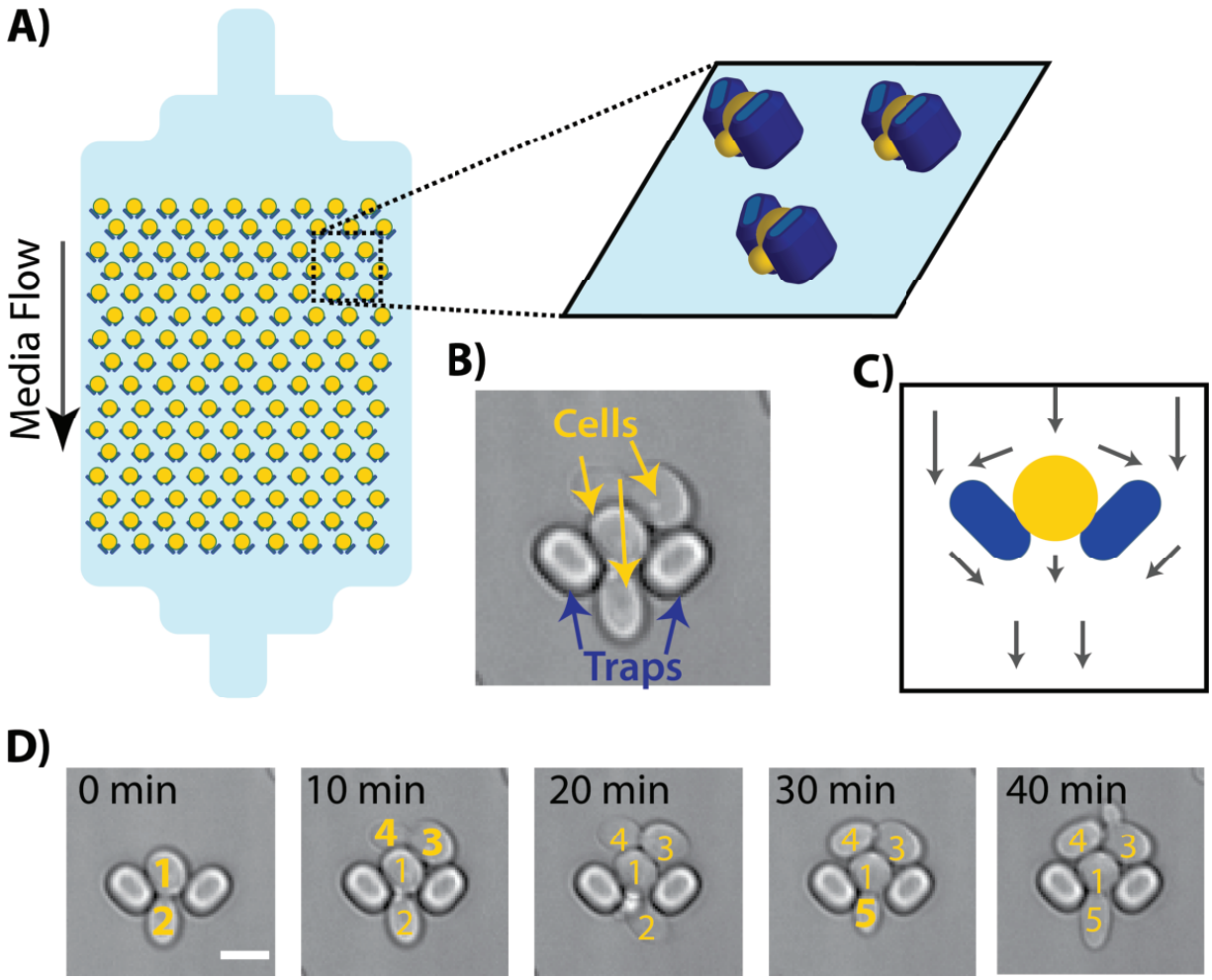

Figure 1: Microfluidic devices for long term imaging of cells impose physical constraints. A) A microfluidic device for budding yeast where cells (yellow objects) are pushed against traps (blue objects) by the flow of media and imaged over long periods. B) An image of a single trap containing multiple cells. The cells and traps share many similarities in shape and optical properties. C) The trap's design imposes specific physical constraints on where cells can be located and where they are likely to move. Arrows represent the flow of fluid. D) Time-lapse images of a single trap show both the appearance of new cells washed in from above and the disappearance of daughters washed away after birth. Cells are individually labeled to show the continuity between time points and the appearance of new cells (bold). Scale bar is 5um.

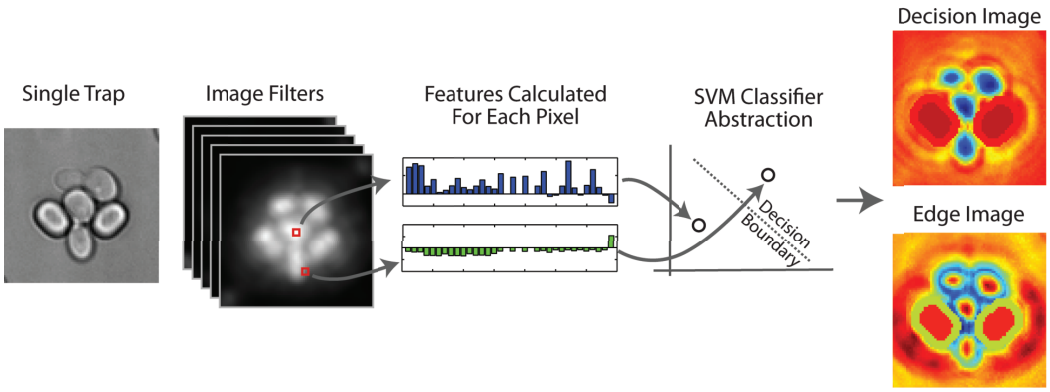

Figure 2: A support vector machine allows the use of almost 60 features to robustly determine probable cell centres. Two bright-field images, one above and one below the plane of focus, are captured at each time point, and a large range of transformations are applied to generate a set of features for each pixel. Two linear SVMs, trained on a curated set of images, provide a probability for the pixel to be either background, in the interior of a cell or on the edge of a cell. These probabilities are reconstituted into a decision image and an edge image, which are used to seed probable cells and identify probable cell edges. In the image, low values (blue) indicate cell interiors in the decision image and cell edges in the edge image.

6 Methods

Identifying physical features of the microfluidic device

Microfluidic devices with traps have floor to ceiling pillars that hold cells and create regular optical features ([31]). Such microfluidic features are not only consistent, stable landmarks, but predictably constrain cellular motion (Fig. 1). We therefore use these physical landmarks at all stages of processing to inform and constrain the segmentation and so increase the accuracy of both segmentation and tracking.

To locate the microfluidic features, the software predicts the locations of traps by performing a normalized cross-correlation ([32]) between the initial time point of the experiment and a canonical image of the microfluidic features. Following this prediction, feedback by the user is required to correct (add or remove) any features that were inaccurately detected. The importance of identifying the locations of the traps mandates input from the user, but this input is only performed at the initial time point and consequently is not laborious. Following this identification, the microfluidic traps are tracked through time to correct for any motion and drift by the stage of the microscope.

Supervised classification of pixels

Supervised classification of pixels has frequently proved to be an efficacious first step in image segmentation ([33, 34, 35, 36, 37]). In these methods, a training set is constructed from images in which pixels are assigned to a discrete set of categories (for example: the image background, the edge of a cell and the interior of a cell). This ground-truth dataset is then used to train a classifier to categorise pixels based on features extracted from the image, which can be either defined beforehand ([36]) or learnt by the classifier ([34, 38]). Such supervised classifiers can leverage multiple features to improve accuracy ([33]) and can often be retrained for different imaging modalities by only changing the training set and not the underlying algorithm ([36, 34]).

In constructing our classifier, we used an approach similar to Sommer et alṪo classify pixels in an image, a large set of predefined transforms are run on the image to extract information about both local properties and properties dependent on the invariant locations of the microfluidic traps. These features are then fed into two linear support vector machines (SVM) classifiers, which have been trained on ground-truth pixels. The first SVM is trained on all pixels to classify pixels as either foreground or background; the second SVM is trained solely on foreground objects to separate these objects into cell

edges and interiors. This two-stage structure was chosen because the foreground and background are relatively easy to discriminate and removing background pixels improves the separation of interior and edge pixels. Extracting the features from the pixels is an order of magnitude slower than classification, and we therefore use the same feature set for both stages of the classification.

Training an SVM corresponds to finding a decision boundary in the space of features that separates the two categories in the training set. To determine this decision boundary, we define the training set as the set of vector-label pairs $\{\mathbf{x}_1, y_1), ..., (\mathbf{x}_N, y_N)\}$ where $\mathbf{x}_i$ is a point in the feature space, y_i indicates the class $\mathbf{x}_i$ belongs to and λ is a slack variable that constrains the cost of misclassification. The support vector, $(\mathbf{w}, b)$, where $\mathbf{w}$ is the vector containing the weight for each feature used in the classification and b is an offset, is selected to minimize

$$\lambda \|\vec{\mathbf{w}}\|^2 + \sum_{i=1}^{N} \max[0, (1 - y_i g(\mathbf{x}))]$$

with the score of each pixel being its distance from the decision boundary:

$$g(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} - \mathbf{b}$$

For a given image to be segmented, the features are calculated and passed to the SVMs to generate two scores for each pixel:

- $g_{\mathrm{BG}}(\vec{\mathbf{x}})$: the score from the SVM for the foreground vs. background
- $g_{\mathrm{edge}}(\vec{\mathbf{x}})$: the score from the SVM for the edge vs. centre

Interpreting these scores as Bayes factors determining the ratios of the probabilities for the identity of each pixel, we can write:

$$g_{\mathrm{BG}} = \log\left(\frac{p_{\mathrm{background}}}{p_{\mathrm{interior}} + p_{\mathrm{edge}}}\right) \tag{1}$$

where $p_{\mathrm{background}} + p_{\mathrm{interior}} + p_{\mathrm{edge}} = 1$ and

$$g_{\mathrm{edge}} = \log\left(\frac{p_{\mathrm{edge|foreground}}}{p_{\mathrm{interior|foreground}}}\right) \tag{2}$$

with $p_{\mathrm{foreground}} = p_{\mathrm{interior}} + p_{\mathrm{edge}}$. Using the product rule for probabilities, we can solve these equations to find $p_{\mathrm{background}}$, p_{interior}, and p_{edge} in terms of g_{BG} and g_{edge} and so calculate the two Bayes factors of interest (see SOM for more details):

$$\text{interior Bayes factor} = \log\left(\frac{p_{\mathrm{background}} + p_{\mathrm{edge}}}{p_{\mathrm{interior}}}\right) \tag{3}$$

$$\text{edge Bayes factor} = \log\left(\frac{p_{\mathrm{background}} + p_{\mathrm{interior}}}{p_{\mathrm{edge}}}\right). \tag{4}$$

The interior Bayes factors for all pixels are reconstituted into a new image of the same size as the original, where low values indicate pixels likely to be cell interiors (Fig. 2). We refer to this image as the **decision image**. The edge Bayes factors for all pixels are similarly reconstituted into what we call the **edge image**, where low values indicate pixels likely to be on a cell edge. These images are used for identifying cell seeds and edges.

For the images we acquired, we used out of focus bright-field images ([17]), which provided the most consistent imaging conditions. Although differential interference contrast (DIC) does give high contrast images, the gradient is dependent on the orientation relative to the centre of the cell, which complicates segmentation ([39]). We acquired images both 2 um above and below the central focal plane to give two distinct bright-field image (either with bright cells and dark edges or dark cells and bright edges). Both images were used in the generation of features for classifying pixels.

When applying the algorithm to the phase contrast images from [23], we retrained the classifier for these images, but did not change any of the transforms used to generate features. The range of features is therefore broad enough that DISCO can be used for diverse images without altering the underlying code and algorithm.

A complete list of all 56 features is in SOM, but we include the radial Hough transform, image smoothing and sharpening features, and relational features to incorporate proximity to the microfluidic traps. Ranking and importance of the features is shown in the SOM and is determined by the weights associated with each feature from the training of the SVM. Training and prediction were performed using the publicly available `liblinear` library ([40]). Both polynomial and RBF-kernel SVMs (using the `libSVM` library) were tested, but offered negligible improvements in accuracy despite a large increase in run-time. We use a five-fold cross-validation approach to determine the cost parameter, λ.

Segmentation using a morphologically constrained model of cell shape

Cells often have a constrained morphology, and including this information can dramatically improve the accuracy of segmentation ([41, 42, 43, 44, 45, 46]) . Here we employ a model of cell shape based on budding yeast, which typically have round to elliptical morphologies. Although the shapes of young cells are constrained, these constraints become less rigid as cells age and become more irregular.

Methods based on active contours provide a straightforward and physically motivated means of encoding information on shapes and have been used extensively for image segmentation ([47, 48, 49, 35]), including for *S. cerevisiae* ([50, 51]). The boundary of a cell is defined by a deformable contour parameterised by a small number of parameters ([47, 52]). The image to be segmented is processed to give a forcing image in which pixels that are likely to be part of an edge have low values. The 'best' contour is then found by minimising a cost function that depends on both this forcing image and the shape of the contour. If the same object is seen in multiple frames of a time-lapse movie, the cost function can also include terms spanning time points to punish physically improbable changes in the object's outline. Further, if the interior pixels of the object can be highlighted, a region term can be added to enforce inclusion of an object's interior within its boundary ([53]). Such methods can additionally integrate prior knowledge on the range of possible shapes with the image data.

Given that cells of *S. cerevisiae* generally have ovoid, concave shapes, we use the two-dimensional polar coordinates, r and θ, to define our contour. If s is a periodic cubic B spline with six evenly spaced knots at fixed angles in the range 0 to 2π and denoted by $\mathbf{r}$ (a vector with six elements), then the contour is all pixels intersected by the curve (Fig. 3Aii):

$$r = s(\theta, \mathbf{r}). \tag{5}$$

This definition allows a range of physically reasonable cell shapes with only six parameters (the elements of $\mathbf{r}$) and balances the competing interests of complexity and flexibility ([49]).

We use both a forcing image, the edge image, and a region term, the decision image. Following others ([53, 50]), we also add an inflation term to prevent the contour collapsing onto a single pixel of low value. These three terms together make the first component of our cost function:

$$\begin{aligned} F\Big(s(\theta, \mathbf{r}_t)\Big) &= \frac{\sum_{p \in \text{pixels on } s} \text{edge image}(p)}{N_{\text{pixels on } s}} \\ &+ \frac{\sum_{p \in \text{pixels within } s} \text{decision image}(p)}{N_{\text{pixels within } s}} \\ &+ \frac{c}{N_{\text{pixels within } s}} \end{aligned} \tag{6}$$

where c is a tuning parameter. If the edge and interior pixels have been accurately identified by the classifier, this cost function will maintain edge pixels on its boundary and interior pixels in its interior and have lower values for larger cells.

We next include terms in our cost function that push the contour towards physically reasonable shapes. To do so, we used a dataset of manually curated cell shapes to determine the empirical distribution of the parameters defining the morphological space.

For new cells, which are not in any previous images, we fit a multivariate normal distribution for the parameter vector $\mathbf{r}$, which is added to the cost function to punish unphysical morphologies. With F as the forcing term (Eq. 6) and $\mathcal{N}$ as the probability density of the normal distribution with parameters μ and Σ fitted to the curated data, the cost function becomes

$$C_{\text{new cells}}(\mathbf{r}_t) = F(\theta, \mathbf{r}_t) - \alpha \log\left[\mathcal{N}(\mathbf{r}_t, \mu, \Sigma)\right] \tag{7}$$

where α is a tuning parameter. We applied the Jarque-Bera test ([54]) to confirm a normal distribution appropriately modelled the data (SOM).

For tracked cells, we find that it is advantageous to include cell growth. Cells are more likely to grow than shrink, although often keep the same shape. To capture these growth effects, we use a log-normal distribution, which has a positive skewness. Fitting the distribution to the element-wise division of the parameter vector for the cell at the current time point ($\mathbf{r}_t$) by the parameter vector for the same cell at the previous time point ($\mathbf{r}_{t-1}$) punishes the relative change in shape, rather than the absolute change, which improves identifying the outline of larger cells. We curated a time-lapse dataset to fit the multivariate log-normal distribution to the element-wise division. Writing $\ln\mathcal{N}$ as the probability density of the log-normal distribution with parameters μ' and Σ' fitted to the curated data and the vector $(r_t^{(1)}/r_{r-1}^{(1)}, \cdots, r_t^{(6)}/r_{r-1}^{(6)})$ as $\mathbf{r}_{t/t-1}$, the final cost function is:

$$C_{\text{tracked cells}}(\mathbf{r}_t) = F(\theta, \mathbf{r}_t) - \beta \log\Big[\ln\mathcal{N}\left(\mathbf{r}_{t/t-1}, \mu', \Sigma'\right)\Big] \tag{8}$$

with β being another tuning parameter.

For both new and tracked cells, the boundary of the cell is found by optimising this cost function for $\mathbf{r}_t$ using a Powell- like line-search algorithm ([55, 56], Fig. 3B).

Incorporating temporal information to refine the prediction of cell centres

Yeast grow and divide, and coupling information on temporal tracking with knowledge on fluid flow can increase the accuracy of identifying cells. Fluid flow on the small length scales of microfluidic devices has a low Reynolds number and so is predictable and consistent. The cell traps and predictable flow profile affect both where cells are initially located and where cells are likely to move to as they grow.

For time points after the first, we developed a method that incorporates this prior knowledge. We generate a prior image $m(x, y)$ for the motion of each cell at the previous time point, which encodes the probability that the centre of the cell has moved to the point (x, y) at the current time point. The motion prior is indexed by a cell's size and location in the trap and is generated from empirical measurements: we use the displacement vectors between the curated pairs of cells used to train the distribution of tracked cell shapes (Eq. 8). When a motion prior is calculated for a particular cell, two probability densities are retrieved, one indexed by its size and one indexed by its location, and the average returned as the motion prior for that particular cell.

To combine this motion prior with the likely cell locations at the current time point, a **probable location image** is calculated for each cell as:

$$\text{probable location image}_i(x, y) = \log[m_i(x, y)] - \text{decision image}(x, y). \tag{9}$$

Fig. 4 shows some probable location images.

Greedy optimization of cell contours

To select cell seeds we use a heuristic based on greedy optimization, encoding the assumption that cells centred on good pixels in the decision are more likely to be accurate.

For the first time point, the decision image is calculated and the first seed is selected as the pixel in the decision image with the lowest value (i.e. the pixel most likely to be a cell centre). Provided the value of this pixel is below a user-defined threshold for identifying new cells, the active contour algorithm is applied using the cost function of Eq. 7 and returns the outline of the putative cell and its score: the value of the cost function. If the score is below another user-defined threshold, the cell

is deemed to be a true cell and assigned a unique label for tracking. The cell is then blotted out of the decision image so that no new seeds will be found within previously identified cells. The procedure is repeated until no pixels remain that are below the user's threshold.

At subsequent time points, the set of probable location images is used to generate cell seeds. We use a similar greedy optimisation: identifying the highest scoring pixel in the set of images, applying the active contour algorithm with the cost function of Eq. 8 centred on this highest scoring pixel, and storing the cell if its score and change in shape meet appropriate thresholds. With each successful identification, the probable location and decision images are modified to include the newly identified cell. Once all previously identified cells have been tracked, new cells are identified by applying the iterative procedure for the first time point to the modified decision image.

This division into tracked cells and new cells has a number of advantages: improving consistency in the location and shape of the cells across time by using information on the shape and location of cells at previous time points to identify and segment cells at the current time point; reducing false positive identifications of new cells and false negatives for cells present over multiple time points by allowing both a more lenient criteria to be applied to cells identified at the previous time point and a more stringent one to be applied to new cells; preventing large and irregularly shaped older cells from being confused for multiple smaller new cells by allowing us to delineate the new-cell and existing cell shape models.

The procedure is applied iteratively over all time points, segmenting the time-lapse images and tracking the cells (Fig. 4). Further details, with pseudo code and a comparison with selecting cell seeds without using historical information, is in the SOM.

All aspects of the pixel classifiers and shape information are obtained from the curated ground-truth images, and the software can therefore be retrained for new imaging modalities and microfluidic devices by curating a new set of ground-truth examples. For example, statistics characterising performance for images provided by [23] were acquired using only a retrained classifier. Once trained, DISCO can be used for different datasets from the same microscope with minimal tuning. We characterised DISCO's performance using the same values for the classifier and tuning parameters, demonstrating that modifications for new datasets are not necessary.

7 Comparison with existing methods

To measure performance, we use metrics taken from the ISBI cell-tracking challenge ([57]), and the ground truth for comparison was generated by manually curating multiple datasets. The ISBI approach provides a single score each for tracking and segmentation, which makes it possible to compare methods directly. To ensure the ground-truth datasets were representative of the variability seen in real experiments, the dataset was composed of images acquired over multiple months in different conditions and by different individuals (SOM). To maintain independence and prevent biasing of the measures of performance, we used strains in which a fluorescent reporter was strongly expressed in the cytoplasm. This fluorescent reporter was used for segmentation by applying a circular Hough transform and an active contour fit using the Chan-Vese algorithm, ensuring the initial segmentation was independent of any method to be tested. Following segmentation of the fluorescence channel, we manually curated all outlines using bright-field images. These datasets are separate from those used for training DISCO's pixel classifier. The curation is in two parts: one for quantifying errors in cell shape and size (>1,000 manually curated cell outlines) and a second for errors between time points (>1,200 curated cell trajectories).

The metric for the accuracy of segmentation is the Jaccard index:

$$SEG = \frac{|A \cap B|}{|A \cup B|} \tag{10}$$

where A is the ground-truth outline and B is the outline identified by the algorithm. The inclusion of the intersection and union means that this score punishes over-segmentation and under-segmentation to the same degree. Cells that are either false positives or false negatives receive a score of 0; cells that are perfectly segmented receive a score of 1.

For estimating the accuracy of tracking, the score is based on transformations applied to an acyclic oriented graph ([57, 58]). Each node in the graph is a detected cell at a specific time point, and the edges connect cells identified to be the same at different time points. The error for an individual cell is determined by the number of operations that must be performed to make the acyclic oriented graph for the test dataset match the ground truth. Operations are weighted by the time required to perform them manually (i.e. more mouse clicks incur a higher cost). The tracking cost for each individual cell is normalized by the number of time points for which the cell is present to give a score between 0 and 1.

We selected three alternative approaches for comparison: CellX ([59]), CellSerpent ([50]) and CellStar ([18]). We focused on methods that use bright-field images as does DISCO. CellSerpent and Cellstar are similar to our approach: cell seeds are identified based on specific features of the image and are used as a starting point for segmentation. CellSerpent identifies seeds via the circular Hough transform; CellStar separates the image into foreground (cells) and background (non-cells) through intensity. Unlike DISCO, however, a single heuristic feature is used to identify seeds and temporal information is ignored. Both impose cell morphologies: CellSerpent enforces cell shapes by penalising deviations from circularity; CellStar punishes large changes in arc length. CellX, in contrast, is substantially different from DISCO with little constraint on the shape of cells. Although CellX uses a Hough-based seeding, the cell outline is found by a graph-cut algorithm applied to an edge image generated using a membrane profile and the proposed seed. CellX and CellStar providing tracking, but Cell Serpent does not, and so we added a commonly used tracking methodology ([19]): cells are assigned the same labels if the fraction of overlap between cells at adjacent time points is > 0.5. CellStar, CellX and CellSerpent were designed for single layers of cells.

Prior to comparison, we attempted to optimize the performance of all the packages on each test dataset according to the instructions provided, and with input from the authors for CellX. To ensure fair comparison, traps identified as cells were removed before scoring. For CellStar, we followed the guidelines provided to optimize segmentation using 30 cells for brightfield training, and 22 cells for phase contrast training. Furthermore, the results of the segmentation for CellX, CellSerpent and CellStar were uniformly dilated and eroded over a range of sizes and scored. We used the best score for each test dataset. We note that DISCO was neither re-trained nor optimized on any of the test datasets, and the results used were unaltered without dilation or erosion. The run-time of each software is similar (SOM).

Although DISCO outperforms the alternatives (Fig. 5), we emphasize that these comparisons are to demonstrate the advantages of including *a priori* knowledge and are not a general critique because the other algorithms are designed for vertically constrained cells and images without microfluidic traps. The slight freedom of movement in z makes a consistent membrane profile difficult to define, complicating detecting cell centres and increasing the importance of the temporal and morphology information we incorporate. Additionally, the pillars of the microfluidic traps change the refractive index that many automated focus systems rely on to maintain focus. This can cause small changes in focus to occur sporadically during long timelapses, complicating both the identification of cell centres and cell membranes. Although the median score for segmentation is similar for the comparison software, DISCO performs with reasonable accuracy on all cells giving a higher mean score. This difference is important as the accuracy of the subsequent tracking is dependent upon the segmentation, and indeed DISCO substantially improves upon the alternatives.

To demonstrate DISCO's flexibility, we obtained also images from [23]. These images are from a microfluidic device with an alternative trap design and were obtained with phase contrast rather than bright-field microscopy (Fig. 5B). Of the three datasets obtained, two were used for training and the third for assessing performance. A subset of the test data was curated by hand because no fluorescence channel was available for an unbiased segmentation. For this data too (Fig. 5 C-Phase), DISCO performs well despite its difference from the type of data for which DISCO was developed.

8 Discussion

With the increasing popularity of microfluidic methods, segmenting and tracking cells from images obtained by time-lapse microscopy is a pressing problem that must be addressed to enable high-throughput, high-content image cytometry. We have presented a new approach that improves accuracy by using information on the physical constraints imposed by the design of the microfluidic system, on the expected morphologies of cells, and on cells at earlier time points.

Our algorithm is innovative for combining multi-feature, multi-class supervised classification, trained prior knowledge and temporal information to achieve robust segmentation and tracking. Using multiple features, not only improves accuracy but also reduces the degree of intervention by the user. This increased accuracy in identifying cell seeds then boosts accuracy in segmenting cells, which in turn boosts accuracy in tracking, and can remove the need to adjust for new experiments. Using the prior knowledge available about the imaging system and the cells being studied further improves performance. Although shape- and model-based segmentation are commonly used, our algorithm fits generative models to curated datasets and exploits temporal information, improving robustness and reducing sensitivity to values of parameters. Indeed, the segmentation is independent of imaging modality, relying solely on the probabilities for classifying cell edges and centres. We have shown that in total this design allows DISCO to be retrained for alternative images and trap designs with only minor tuning.

Microfluidic devices can both provide unprecedented quantities of high-content data and enable investigation of the effects of dynamic cellular environments. For ageing research in particular, where traditional approaches are notoriously low throughput ([60]), microfluidic techniques are potentially revolutionary, but only if quantifying the resulting data can be automated. The importance of using more natural, changing environments, such as those that can be generated by microfluidic technologies, to study cellular behaviour is now being recognized ([61, 62]), but again analysing the time-lapse data produced can be a bottleneck to progress ([63]). Although moving to these assays may be as revealing as the switch from studying bulk populations to studying behaviour in single cells, automation of segmenting and tracking cells is necessary. To impact fully, then, microfluidics-based time-lapse microscopy depends on the capability to robustly and consistently process the information in the images acquired using analytical frameworks such as the one we present here.

Acknowledgement

We would like to thank members of the Swain lab for constructive comments, Dr. Lucia Bandiera for providing some of the images used in the comparisons, and Prof. Nan Hao and Dr. Yang Li of U.C. San Diego for providing additional images.

Funding: This work was supported by the BBSRC (MMC & PSS), the Human Frontiers Science Program (MMC, EB & PSS), and the Scottish Universities Life Sciences Alliance (EB & PSS).

References

[1] E. Meijering, A. E. Carpenter, H. Peng, F. A. Hamprecht, J.-C. Olivo-Marin, *Nature Biotechnology* **34**, 1250 (2016).

[2] E. Meijering, *IEEE Signal Processing Magazine* **29**, 140 (2012).

[3] L. Kamentsky, *et al.*, *Bioinformatics* **27**, 1179 (2011).

[4] C. Sommer, D. W. Gerlich, *Journal of Cell Science* **126**, 5529 (2013).

[5] M. Schiegg, *et al.*, *Bioinformatics* **31**, 948 (2015).

[6] F. Federici, L. Dupuy, L. Laplaze, M. Heisler, J. Haseloff, *Nature Methods* **9**, 483 (2012).

[7] Q. Zhong, A. G. Busetto, J. P. Fededa, J. M. Buhmann, D. W. Gerlich, *Nature Methods* **9**, 711 (2012).

[8] C. Conrad, *et al.*, *Nature Methods* **8**, 246 (2011).

[9] M. Held, *et al.*, *Nature Methods* **7**, 747 (2010).

[10] S. Pelet, R. Dechant, S. S. Lee, F. van Drogen, M. Peter, *Integrative Biology* **4**, 1274 (2012).

[11] W. Huh, J. Falvo, L. Gerke, A. Carroll, *Nature* (2003).

[12] Y. T. Chong, *et al.*, *Cell* **161**, 1413 (2015).

[13] T. Zimmermann, J. Rietdorf, R. Pepperkok, *FEBS Letters* **546**, 87 (2003).

[14] V. Magidson, A. Khodjakov, *Methods in cell biology* **114** (2013).

[15] M. R. Bennett, J. Hasty, *Nature Reviews Genetics* **10**, 628 (2009).

[16] A. Doncic, U. Eser, O. Atay, J. M. Skotheim, *PLoS ONE* **8**, e57970 (2013).

[17] A. Gordon, *et al.*, *Nature Methods* **4**, 175 (2007).

[18] C. Versari, *et al.*, *Journal of The Royal Society Interface* **14**, 20160705 (2017).

[19] S.-C. Chen, T. Zhao, G. J. Gordon, R. F. Murphy, *Computational Intelligence and Bioinformatics and Computational Biology, 2006. CIBCB'06. 2006 IEEE Symposium on* (IEEE, 2006), pp. 1–8.

[20] M. Zhan, *et al.*, *PLoS Computational Biology* **11**, e1004194 (2015).

[21] M. M. Crane, *et al.*, *Nature Methods* **9**, 977 (2012).

[22] M. M. Crane, I. B. N. Clark, E. Bakker, S. Smith, P. S. Swain, *PLoS ONE* **9**, e100042 (2014).

[23] Y. Li, *et al.*, *bioRxiv* (2017).

[24] J. Ryley, O. M. Pereira-Smith, *Yeast* **23**, 1065 (2006).

[25] S. Sik, I. Avalos, D. H. E. W. Huberts, L. P. Lee, M. Heinemann, *PNAS* **109**, 4916 (2012).

[26] Y. Zhang, *et al.*, *PLoS One* **7**, e48275 (2012).

[27] Z. Xie, *et al.*, *Aging Cell* **11**, 599 (2012).

[28] M. C. Jo, W. Liu, L. Gu, W. Dang, L. Qin, *PNAS* **112**, 9364 (2015).

[29] K. Li, T. Kanade, *Proceedings of 2nd Workshop on Microsopic Image Analysis with Applications in Biology, Piscataway, NJ, USA* pp. 1–8 (2007).

[30] A. A. Granados, *et al.*, *eLife* **6**, e21415 (2017).

[31] K. L. Chen, M. M. Crane, M. Kaeberlein, *Mechanisms of ageing and development* **161**, 262 (2017).

[32] M. Guizar-Sicairos, S. T. Thurman, J. R. Fienup, *Optics Letters* **33**, 156 (2008).

[33] C. Zhang, F. Huber, M. Knop, F. A. Hamprecht, *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on* (IEEE, 2014), pp. 1267–1270.

[34] D. Ciresan, A. Giusti, L. M. Gambardella, J. Schmidhuber, *Advances in neural information processing systems* pp. 2843–2851 (2012).

[35] F. Xing, S. Member, L. Yang, *IEEE Reviews in Biomedical Engineering* pp. 234–263 (2016).

[36] C. Sommer, C. Straehle, U. Kothe, F. a. Hamprecht, *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, no. 1 (IEEE, 2011), pp. 230–233.

[37] B. T. Grys, *et al.*, *The Journal of Cell Biology* (2016).

[38] Y. Zhou, H. Chang, K. E. Barner, B. Parvin, *Proceedings - International Symposium on Biomedical Imaging* **2015-July**, 1284 (2015).

[39] F. Ning, *et al.*, *IEEE Transactions on Image Processing* **14**, 1360 (2005).

[40] R. en Fan, K. wei Chang, C. jui Hsieh, X. rui Wang, C. jen Lin, *Journal of Machine Learning Research* **9**, 1871 (2008).

[41] R. Delgado-Gonzalo, D. Schmitter, V. Uhlmann, M. Unser, *IEEE Transactions on Image Processing* **24**, 3915 (2015).

[42] M. E. Leventon, W. E. L. Grimson, O. Faugeras, *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)* (2000), vol. 1, pp. 316–323 vol.1.

[43] A. Foulonneau, P. Charbonnier, F. Heitz, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**, 1352 (2006).

[44] Y. Chen, *et al.*, *International Journal of Computer Vision* **50**, 315 (2002).

[45] M. A. Charmi, S. Derrode, F. Ghorbel, *Pattern Recognition Letters* **29**, 897 (2008).

[46] S. Ali, R. Veltri, J. I. Epstein, C. Christudass, A. Madabhushi, *Computerized medical imaging and graphics* **41**, 3 (2015).

[47] T. McInerney, T. McInerney, D. Terzopoulos, D. Terzopoulos, *Medical image analysis* **1**, 91 (1996).

[48] E. C. Garner, *Molecular microbiology* **80**, 577 (2011).

[49] R. Delgado-Gonzalo, M. Unser, *IRBM* **34**, 235 (2013).

[50] K. Bredies, H. Wolinski, *Computing and Visualization in Science* **14**, 341 (2011).

[51] M. Kvarnström, K. Logg, A. Diez, K. Bodvard, M. Käll, *Optics express* **16**, 12943 (2008).

[52] A. Blake, M. Isard, *Active Contours: The Application of Techniques from Graphics, Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion* (Springer Science & Business Media, 2012).

[53] Y. Huang, Z. Liu, *Computational and Mathematical Methods in Medicine* **2015**, 1 (2015).

[54] C. M. Jarque, A. K. Bera, *Economics Letters* **6**, 255 (1980).

[55] M. J. D. Powell, *The Computer Journal* **7**, 155 (1964).

[56] E. Ziegel, W. Press, B. Flannery, S. Teukolsky, W. Vetterling, *Technometrics* **29**, 501 (1987).

[57] M. Maška, *et al.*, *Bioinformatics* **30**, 1609 (2014).

[58] P. Matula, *et al.*, *PLoS One* **10**, e0144959 (2015).

[59] S. Dimopoulos, C. E. Mayer, F. Rudolf, J. Stelling, *Bioinformatics* **30**, 2644 (2014).

[60] M. Kaeberlein, *Nature* **464**, 513 (2010).

[61] P. Nurse, *Nature* **454**, 424 (2008).

[62] R. P. Alexander, P. M. Kim, T. Emonet, M. B. Gerstein, *Science signaling* **2**, pe44 (2009).

[63] J. C. Locke, M. B. Elowitz, *Nature Reviews Microbiology* **7**, 383 (2009).

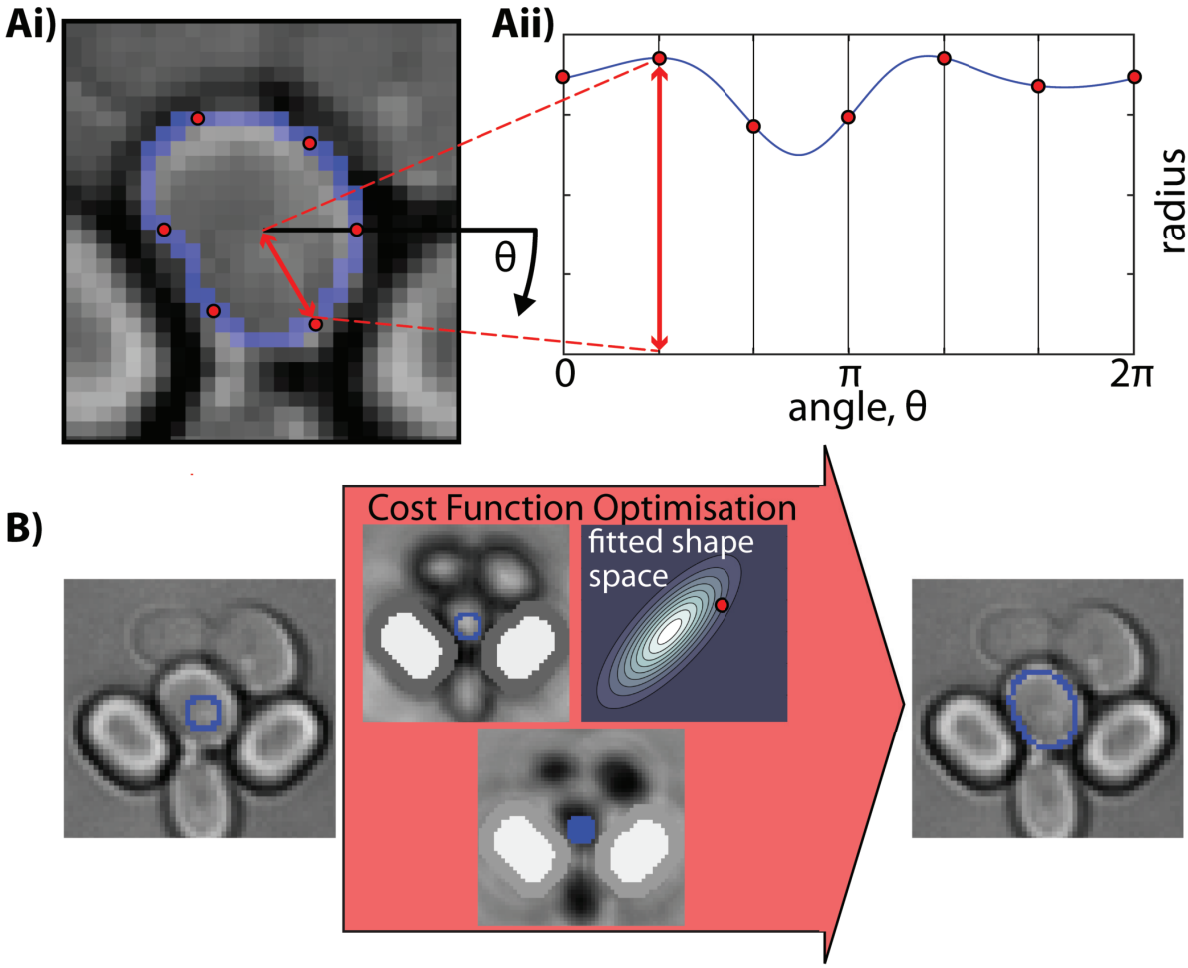

Figure 3: Using a morphological model for cells allows prior information about their shape to be exploited during segmentation. A) The definition of a cell's contour: the contour is defined by a periodic cubic B spline in polar coordinates (blue line in Aii) centred on the cell's seed and is completely defined by the six knots of the spline (red circles in Aii). This spline is mapped back to the coordinates in the image to produce a cell's contour (Ai). B) Identification: we define a cost function combining the values of the pixels in the edge image along the contour, the values of the pixels in the decision image inside the contour and a fitted probability distribution of potential cell shapes (B, centre). Starting from an initial guess (B,left), this cost function is minimised to give the correct contour (B,right).

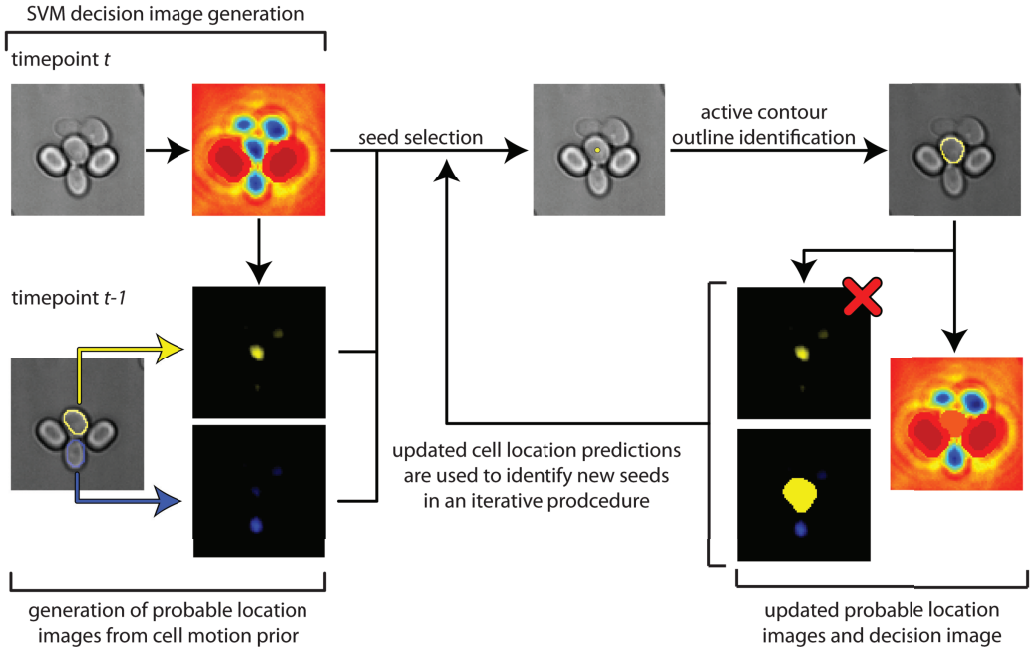

Figure 4: Greedy optimisation prioritises cells identified at previous time points to ensure reliable tracking. To segment cells at the current time point (t), a decision image and probable location images are generated and subjected to greedy optimisation. The probable location images are used to propose a seed: the highest scoring pixel in the image set. If this score is above a threshold, the seed is used to generate a cell contour following Fig. 3. If the contour meets criteria for its shape and its overall score, the contour is stored and the decision and probable location images updated to prevent new cells being found in the contour's location. If the contour is a tracked cell rather than a new cell, the probable location image for this cell is no longer considered when seeds are generated. The procedure is repeated until no pixels remain above threshold. Then, seeds for new cells are proposed from the decision image.

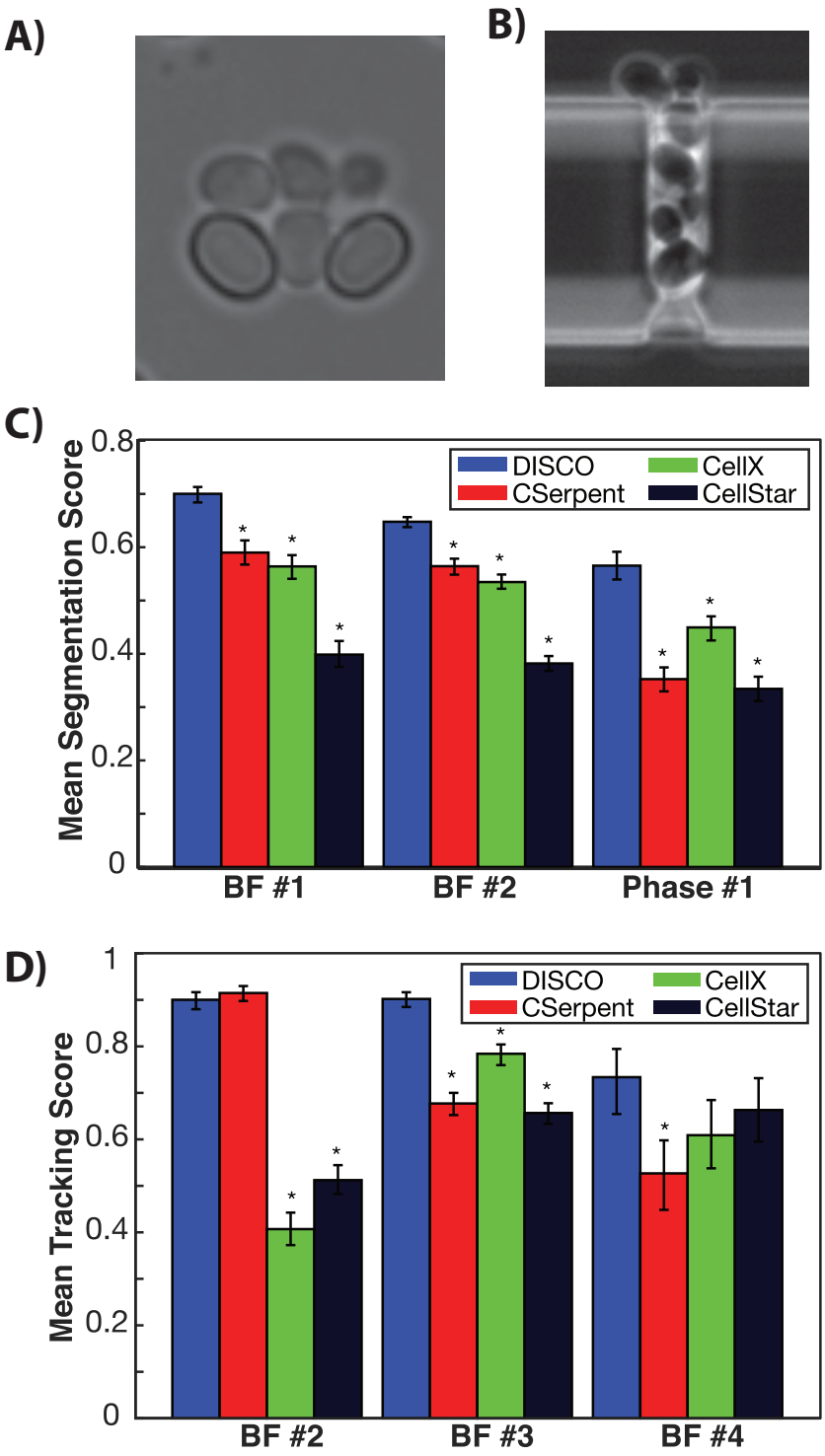

Figure 5: Comparing performance for segmentation and tracking for DISCO, CellSerpent, CellX and CellStar. A) A representative image of cells acquired using brightfield imaging in the ALCATRAS device and which compose the datasets BF 1-4. B) Representative image of cells from [23] which were acquired using phase contrast imaging and compose the Phase dataset. C-D) Each of the four software packages were applied to all cells in the segmentation and tracking ground-truth data sets. Mean performance over all cells in each data set is shown. Error bars are 95% confidence intervals for the mean calculated from bootstrapping with replacement. Differences between software packages are therefore at a 5% significance level when the error bars do not overlap (*). C) Mean segmentation accuracy. DISCO improves on segmentation accuracy on all datasets. D) Tracking: DISCO, CellX and CellStar were run using their native tracking algorithms and CellSerpent was run using overlap tracking. The substantial differences in performance illustrates the challenge these images pose for segmentation.